

ANALISIS SINYAL ELECTRONIC NOSE BERBASIS WAVELET MENGGUNAKAN SUPPORT VECTOR MACHINE UNTUK IDENTIFIKASI JENIS TEH HITAM

Marlus Eri Sitohang

Abstrak, Penelitian ini bertujuan untuk melakukan identifikasi terhadap aroma teh hitam. Aroma teh hitam didapat dari hasil pengukuran alat Electronic Nose (E-nose) yang menggunakan rangkaian sensor gas. Identifikasi aroma teh hitam menggunakan dilakukan dengan Principal Component AnalysisSupport (PCA) dan Support Vector Machine (SMV).

Penelitian ini dilakukan dengan tahapan analisis data keluaran dari rangkaian sensor yang berupa tegangan dengan menggunakan wavelet. Data tegangan listrik berasal dari 4 buah sensor TGS880, TGS826, TGS822, TGS825. Denoising dan kompres dari sinyal keluaran sensor ditransformasi menggunakan wavelet Daubechies yaitu D4, D8 dan D20, dengan tingkat dekomposisi 3 dan 6.

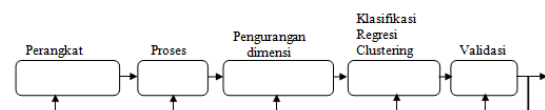
Transformasi PCA dilakukan terhadap fitur ekstrasi yang dihasilkan dari analisis wavelet, sebelum dimasukan ketahap klasifikasi pola. PCA digunakan untuk mengurangi dimensi dari data fitur ekstrasi tanpa kehilangan banyak informasi signifikan yang ada di dalamnya. Hasil pengujian menunjukan SVM mampu memberikan klasifikasi dengan akurasi 97,53% menggunakan wavelet Daubechies D20 dengan tingkat dekomposisi 6.

Index Term - Electronic Nose, Wavelet, Principal Component AnalysisSupport (PCA), Support Vector Machine (SVM)

I. PENDAHULUAN

One Electronic Nose (*e-nose*) adalah alat yang mengkombinasikan rangkaian sensor gas dan analisi pola untuk pedekteksian, identifikasi atau kuantisasi senyawa *volatile*(Gutierrez, 2002). Persaud & Dodd (1982) meneliti tentang pengembangan teknologi *e-nose*, yang dapat mendeteksi dan mengenali aroma bau. Prinsip dasar kerja *e-nose* adalah menirukan fungsi hidung manusia., yang mana di dalamnya dijumpai berbagai reseptor yang dapat mengidentifikasi aroma bau. Perangkat ini telah mengalami banyak pengembangan dan kini digunakan untuk memenuhi kebutuhan industri.

Pada dekate ini, banyak peneliti yang telah berupaya meningkatkan kinerja *e-nose* dan engimprovisasi penggunaan *e-nose*. Saat ini penggunaan *e-nose* tidak hanya untuk membedakan dan klasifikasi sampel aroma bau tapi juga mengukur unsur dari suatu sampel. Selain evaluasi sensori manusia dan identifikasi analitis, *e-noise* dapat menjadi metode pelengkap dalam pengendalian mutu rempah-rempah dan aroma (Moy dkk., 1994). Proses pada *e-nose* dapat dibagi menjadi empat bagian (Gutierrez, 2002) yaitu proses sinyal, pengurangan dimensi, prediksi dan validasi seperti terlihat pada gambar 1.



Gambar 1 Ilustrai proses *e-nose*

II. TEORI

2.1 Wavelet

Analisis wavelet dapat mengungkapka data dari berbagai aspek yaitu teknik analisis sinyal, kecenderungan, titik breakdown, diskontinuitas, dan

Marlus Eri Sitohang : Magister Teknik Informatika, Fakultas Ilmu Komputer, Universitas Bina Nusantara

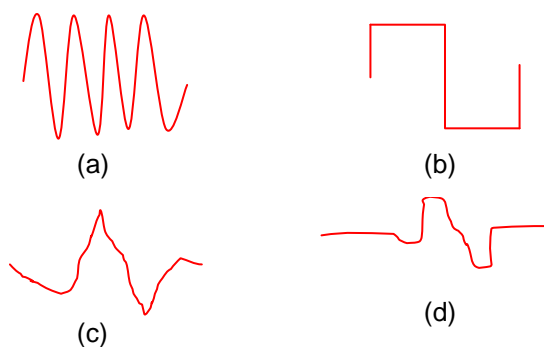
e-mail: (eri321@gmail.com)

Jurnal Sistem Komputer - Vol.2, No.2, November 2012, ISSN: 2087-4685, e-ISSN: 2252-3456

Jsiskom - 47

bentuk kesamaan (Misiti, dkk. 2004). Dikarenakan data dapat direpresentasikan dalam bentuk yang berbeda-beda, wavelet analisis dapat memperkecil sinyal gangguan tanpa adanya degradasi.

Analisis wavelet memiliki berhubungan dengan analisis Fourier. Didalam analisis Fourier, sinusoids pada gambar 2(b) digunakan sebagai basis fungsi dengan variasi frekuensi dan perluasan dari $-\infty$ sampai $+\infty$. Dengan kata lain, kelompok wavelet families have batas seperti yang digambarkan pada gambar 2(b,c,d).



Gambar 2 Illustration basis Fourier basis and basis wavelet

Sebuah sinyal atau fungsi $f(t)$ dari ruang S dapat di analisa dan diproses dari berbagai aplikasi yang berbeda jika bisa direpresentasikan sebagai kombinasi linear (Burrus, 2003), seperti diperlihatkan pada persamaan (2.3).

$$f(t) = \sum_i \alpha_i \varphi_i \quad (1)$$

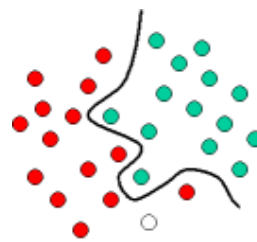
Dimana α adalah koefisien ekspansi, φ adalah fungsi ekspansi t , i adalah index integer

Tahap pertama analisis wavelet adalah menentukan tipe wavelet, yang disebut dengan mother wavelet atau analyzing wavelet, yang akan digunakan. Hal ini perlu dilakukan karena fungsi wavelet sangat bervariasi dan dikelompokkan berdasarkan fungsi dasar wavelet masing-masing. Contoh keluarga wavelet adalah wavelet Daubechies.

Setelah pemilihan mother wavelet, tahap selanjutnya membentuk basis wavelet yang akan digunakan untuk mentransformasikan sinyal. Suatu basis dapat dibentuk dengan mengubah nilai translasi dan dilasi dari mother wavelet-nya. Analisis temporal dilakukan dengan menggunakan basis wavelet frekuensi tinggi, sedangkan analisis frekuensi dilakukan dengan menggunakan basis wavelet frekuensi rendah. Operasi terhadap suatu sinyal hanya dapat dilakukan dengan menggunakan koefisien-koefisien wavelet yang berhubungan.

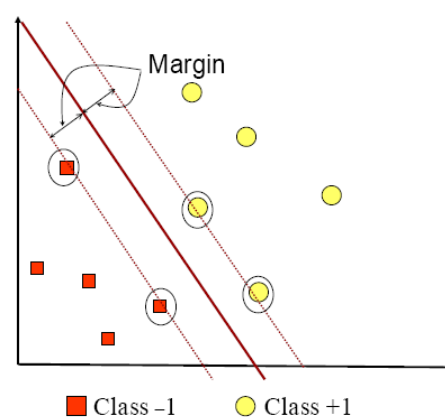
2.2 Support Vector Machine

Support Vector Machine (SVM) merupakan salah satu teknik supervised learning dimana pembelajaran dilakukan dengan menggunakan data-data yang sudah jelas kelasnya (Vapnik). SVM merupakan nama untuk suatu himpunan metode yang dapat digunakan untuk melakukan klasifikasi dan regresi. Dalam penelitian ini, metode SVM yang digunakan adalah metode untuk melakukan klasifikasi. SVM didasarkan pada konsep bidang keputusan. Bidang keputusan berfungsi untuk memisahkan obyek-obyek yang memiliki kelas yang berbeda. Gambar 3 memberikan ilustrasi mengenai satu bidang keputusan yang memisahkan dua kelas, yaitu hijau dan merah.



Gambar 3 Bidang Keputusan yang Memisahkan Dua Kelas.

Ide dasar dari SVM adalah menemukan fungsi *hyperplane* yang mampu memisahkan antara dua kelas dengan optimal. Optimal di sini artinya adalah *hyperplane* tersebut mampu memisahkan kedua kelas dengan margin yang maksimal. Margin adalah jarak antara garis *hyperplane* dengan anggota-anggota terdekat dari kedua kelas. Hal ini diilustrasikan pada gambar 4

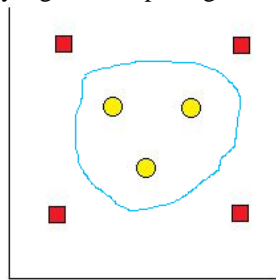


Gambar 4 Hyperplane dengan Margin Maksimal.

2.3 SVM Non-Linear

Dalam penerapannya di lapangan, seringkali

ditemukan kasus-kasus yang tidak dapat dipecahkan oleh SVM linear. Salah satu pemecahan terhadap masalah tersebut adalah penggunaan soft margin namun, dalam beberapa kasus soft margin kurang tepat untuk diterapkan. Suatu model SVM yang lebih dikenal sebagai SVM non-linear kemudian dikembangkan untuk memecahkan masalah ini. Sesuai dengan namanya, SVM ini mampu menghasilkan hyperplane dengan persamaan non-linear, seperti yang terlihat pada gambar 4.



Gambar 4 Contoh Penyebaran Data dan Hyperplane Non-linear.

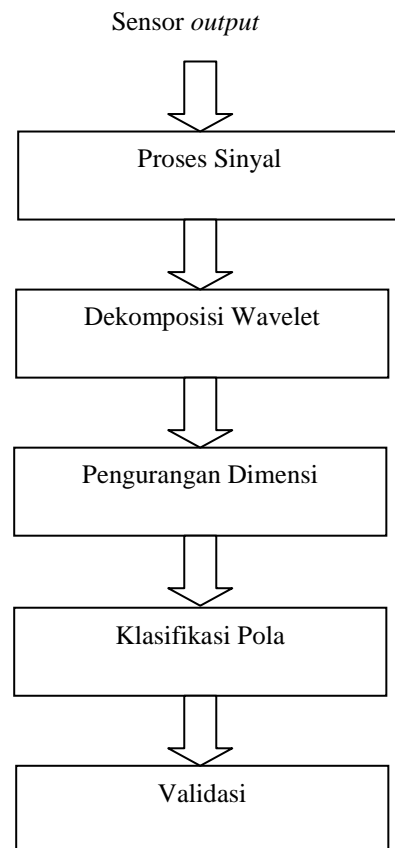
Ide dasar dari model SVM non-linear ini adalah memetakan data dari suatu bidang dengan dimensi tertentu, ke dalam bidang dengan dimensi yang lebih tinggi. Bidang tersebut kemudian dikenal dengan nama bidang input dan bidang fitur. Pemetaan dari bidang input ke bidang fitur (F) ini secara matematika dilambangkan dengan $\Phi: \mathcal{R}^N \rightarrow F$. Data hasil pemetaan tersebut dilambangkan dengan $x \in \Phi(x)$.

2.4 SVM Multi-Kelas

Pada dasarnya, SVM dirancang hanya untuk masalah klasifikasi biner, dimana data yang terdapat di dalam suatu model SVM hanya mungkin terklasifikasi ke dalam dua kelas saja. Dalam kasus dunia nyata, kemampuan untuk melakukan klasifikasi hanya ke dalam dua kelas seringkali tidak cukup karena banyaknya variasi yang mungkin ditemui. Karenanya, digunakan cara tertentu untuk memungkinkan SVM mampu melakukan klasifikasi terhadap lebih dari dua kelas dimana caranya adalah dengan menggunakan lebih dari satu persamaan hyperplane. Hingga saat ini, ada tiga pendekatan yang dapat digunakan untuk memecahkan masalah SVM multikelas ini, yaitu metode *one-against-all*, *one-against-one*, dan *directed acyclic graph*. Pembahasan berikutnya akan membahas lebih dalam mengenai SVM multikelas dengan menggunakan contoh 3 kelas (A, B, dan C) kemungkinan klasifikasi.

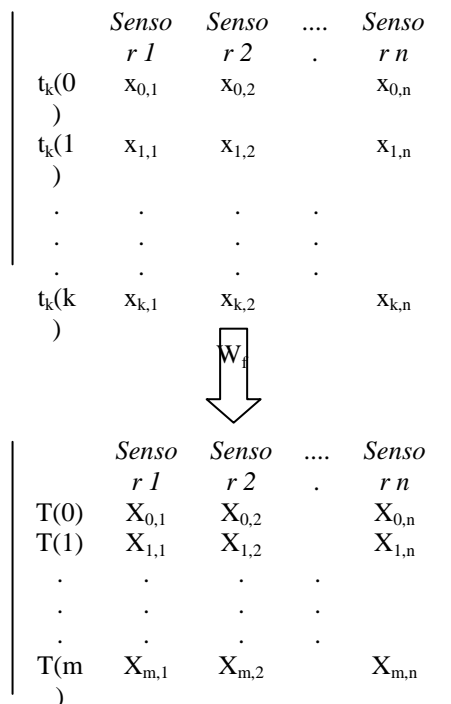
III. METODOLOGI

Alur diagram pada Gambar 5 merupakan ilustrasi dari penelitian yang diajukan.



Gambar 5 Diagram alur penelitian

Sinyal ditransformasikan menggunakan *discrete wavelet transformation*. Perubahan data sinyal dapat direpresentasikan dalam notasi matriks seperti yang disajikan dalam pada Gambar 6



Gambar 6 Perubahan data sinyal

Dimana:

- $m \ll k$
- $x_{0,1}$ adalah respon sinyal dari sensor 1 pada waktu $t=0$
- $X_{0,1}$ adalah koefisien wavelete dari sensor 1 pada waktu $T=0$

Linear feature extractor digunakan untuk menentukan pendekatan *subset*. Pengurangan dimensi menggunakan Principal Component Analysis (PCA) yang telah dikenal juga sebagai Karhunen-Loève.

Pengenalan pola-pola sampel teh hitam pada penelitian ini menggunakan Support Vector Machine (SVM). Dalam proses ini, data hasil pengukuran yang telah di ekstrak menggunakan PCA digunakan sebagai data masukan pada proses SVM. Data tersebut dibagi menjadi dua bagian, yaitu data pelatihan dan data pengujian. Data pelatihan digunakan untuk melatih mesin pembelajaran mengenali pola-pola setiap sampel. Sedangkan data pengujian digunakan untuk menguji kemampuan jaringan dalam mengenali pola-pola sampel.

Metode validasi yang digunakan dalam penentuan data pelatihan dan pengujian adalah ini merujuk pada metode *hold out*. Dalam metode ini, data asli dipartisi menjadi dua himpunan yang saling terpisah yang dinamakan *training set* dan *test set*. Model klasifikasi kemudian dibangun berdasarkan training set dan hasilnya kemudian dievaluasi dengan menggunakan testing set. Akurasi dari masing-masing metode klasifikasi dapat diestimasi

berdasarkan akurasi yang diperoleh dari test set. Proporsi antara *training set* dan *test set* tidak mengikat tetapi agar variansi dalam model tidak terlalu besar maka dapat ditentukan bahwa proporsi training set lebih besar daripada test setnya. Secara khusus, 2/3 dari data dijadikan *training set* dan 1/3 lagi dijadikan *testing set*.

Bahan yang digunakan dalam penelitian ini berupa sampel teh hitam. Data dalam penelitian ini diperoleh dari Laboratorium Penelitian dan Pengujian Terpadu (LPPT) UGM. E-nose yang digunakan dalam penelitian ini menggunakan 4 buah sensor gas, yaitu :

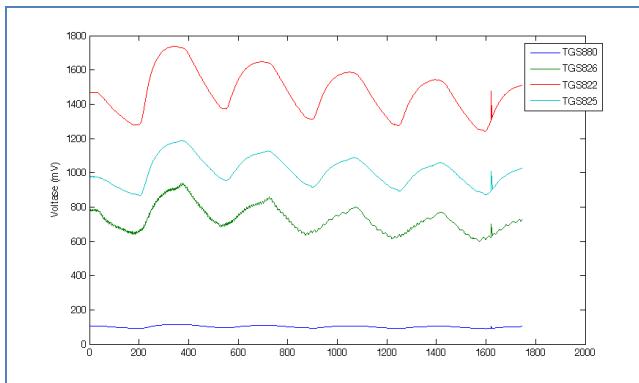
- TGS880 yang peka terhadap bau/aroma secara umum
- TGS826 yang peka terhadap amonia
- TGS822 yang peka terhadap gugus alkohol
- TGS825 yang peka terhadap hidrogen sulfida

3 Analisa

3.1 Aroma Teh

E-nose terdiri dari 4 buah sensor gas dimana masing-masing dari sensor tersebut memiliki kepekaan yang berbeda pada saat pendeteksian aroma. Jenis teh yang diuji dalam penelitian ini merupakan jenis teh hitam yang diperoleh dari Pusat Teh dan Kia (PPTK). Pengujian aroma teh dengan menggunakan E-nose yang dilengkapi dengan 4 sensor kimia, akan memperoleh data perubahan voltase dari masing-masing jenis teh hitam tersebut.

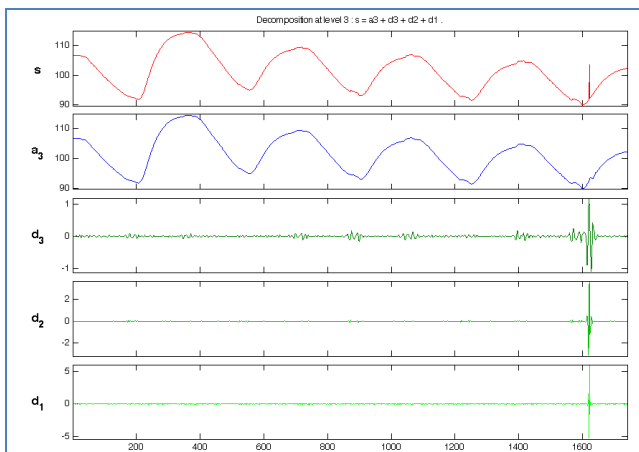
Grafik perubahan voltase pada masing-masing sensor dapat dilihat pada gambar 7. Perubahan voltase terbesar ada pada sensor TGS822, dimana sensor TGS822 memiliki kepekaan terhadap gugus alkohol. Walaupun setiap sensor didesign hanya sensitif terhadap aroma teh tertentu, namun setelah disusun secara kolektif dalam bentuk rangkaian sensor gas, keluaran masing-masing sensor akan memberikan kontribusi dalam membentuk pola unik pada setiap jenis teh (Triyana, 1997). Menurut Yamanishi (1995) komposisi kimia yang mempengaruhi aroma teh berupa senyawa-seyawa *hydrocarbons*, *alcohols*, *aldehydes*, *acids*, *esters*, *sulfur* dan *nitrogenous*.



Gambar 7 Grafik perubahan voltase pada sampel teh hitam jenis BP2

3.2 Ekstraksi Ciri

Analisis wavelet dilakukan dengan menggunakan *toolbox* sinyal wavelet dari perangkat lunak Matlab 7.0.1. Sinyal *E-nose* seperti dapat dilihat pada gambar 8 diuji untuk denoising dan kompresi menggunakan Daubechies db4 dan tranformasi di level 3 dari dekomposisi. Setelah dilakukan dekomposisi level 3, komponen sinyal asli terpadu dalam perkiraan (a_3) dan *noise* terkonsentrasi dalam grafik (d_1 , d_2 , d_3) dengan informasi tentang sinyal. Kompresi dari dekomposisi sinyal *E-nose* dengan kesalahan $1,75 \cdot 10^{-11}$ memungkinkan dengan hanya 13,12% total wavelet koefisien. Hasil analisis menunjukkan bahwa untuk db8 menggunakan dekomposisi level 3, menghasilkan error $4,22 \cdot 10^{-11}$ dengan 13,53% koefisien wavelet. Dengan menggunakan db20 menghasilkan error $2,43 \cdot 10^{-10}$ dengan 14,76 % koefisien wavelet. Selanjutnya, untuk fungsi wavelet yang sama (db4, DB8 dan Db20), tingkat dekomposisi diubah dari 3 menjadi 6. Untuk level dekomposisi 6, sinyal rekonstruksi adalah mungkin dengan hanya 2,00%, 2,41% dan 3,82% dari jumlah total koefisien wavelet.

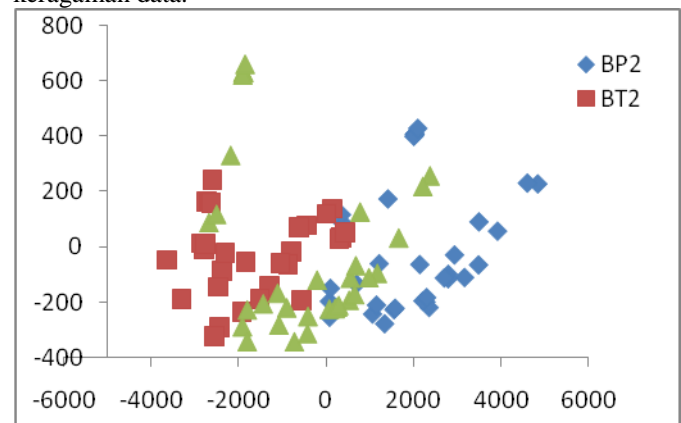


Gambar 8 : Dekomposisi dari sinyal *E-nose* (sensor TGS880 pada teh BP2) level 3 dengan menggunakan Daubechies Db8 wavelet transform

3.3 Pengurangan Dimensi

Analisis Komponen Utama banyak dalam penelitian ini digunakan untuk memproyeksikan atau mengubah kolom data berukuran 4 menjadi bentuk sajian data dengan ukuran 2. Transformasi PCA terhadap sebuah ruang data yang besar akan menghasilkan sejumlah vektor basis ortonormal dalam bentuk kumpulan vektor eigen dari suatu matriks kovarian tertentu yang dapat secara optimal menyajikan distribusi data. Sasaran dari PCA adalah menangkap variasi total dari citra yang ada di dalam basis data yang dilatihkan. Untuk kemudian mereduksinya sehingga menjadi variabel-variabel yang lebih sedikit. Dengan reduksi ini maka waktu komputasi dapat dikurangi dan kompleksitas dari ciri yang tidak perlu dapat dihilangkan.

Gambar 9 memperlihatkan kumpulan data yang sudah di transformasi dengan PCA. Jumlah 4 sensor menggambarkan 4 dimensi data, kemudian di transformasi menjadi dimensi 2 dimensi yaitu PC1 dan PC2. Hasilnya memperlihatkan bahwa Principle Component atau Komponen Utama ke-1 dan ke-2 (PC1 dan PC2) mampu menyerap 97,54% dan 1,88% keragaman data.



Gambar 9 PCA pada semua data teh (BP2, BT2, DT2)

3.4 Klasifikasi pola

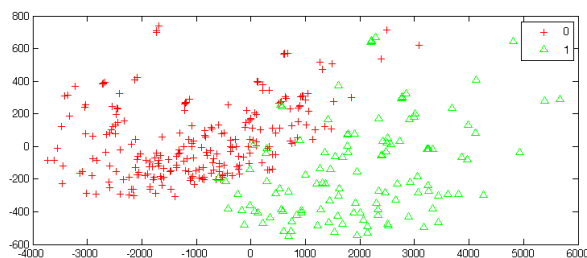
Klasifikasi pola pada penelitian ini menggunakan *Support Vector Machine* (SVM).

Support Vector Machine (SVM) merupakan salah satu teknik supervised learning dimana pembelajaran dilakukan dengan menggunakan data-data yang sudah jelas kelasnya. SVM mampu melakukan klasifikasi dengan baik. Hal ini disebabkan karena pencarian solusi optimal yang dilakukan oleh SVM bertujuan untuk mencari solusi optimal berdasarkan data yang dimiliki

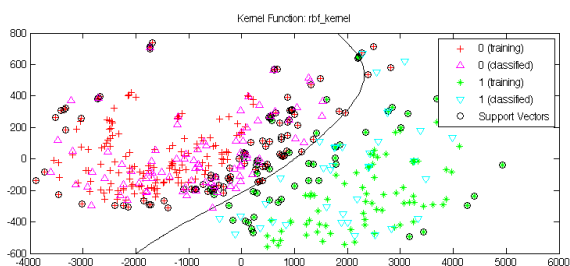
sekarang ini, dan bukan bertujuan mencari solusi yang optimal untuk data yang tidak terbatas.

Pada dasarnya, SVM dirancang hanya untuk masalah klasifikasi biner, dimana data yang terdapat di dalam suatu model SVM hanya mungkin terklasifikasi ke dalam dua kelas saja. Walaupun demikian dilakukan sebuah pendekatan agar dapat dilakukan pemecahan untuk dapat mengklasifikasi lebih dari 2 kelas. Ada 3 cara untuk solusi SMV multi kelas, yaitu metode *one-against-all*, *one-against-one*, dan *directed acyclic grap*. Pada penelitian ini dilakukan SVM mutti kelas dengan menggunakan *one-against-all*, dimana setiap data yang terdapat di dalam salah satu kelas SVM akan dibandingkan dengan gabungan dari data- data yang bukan merupakan anggota kelas tersebut.

Gamabar 10 menunjukan sampel yang akan diuji dengan SVM. Data 1 menggambarkan data dari BP2 dan 0 merupakan data selain data BP2 (data BT2 dan DT2). Gambar 11 merupakan hasil dari pembelajaran menggunakan SVM multi kelas dengan karnel Gaussian. Hasil yang didapat dari pembelajaran tersebut menghasilkan performa klasifikasi sebesar 94,49%



Gamabar 10 Sampel data BP2



Gamabar 11 Klasifikasi data BP2

Hasil klasifikasi data teh hitam ditampilkan pada table 1. Untuk semua data nilai klasifikasi diperoleh dari ekstrasi fitur menggunakan DWT db4, db8 dan db20 pada tingkat dekomposisi 6.

Tabel 1 Hasil klasifikasi Menggunakan SVM dan Neural Network

Teh	Klasifikasi (%)
-----	-----------------

	SVM			NN	
	BP2	BT2	DT2	Neuro n	Klasifikasi (%)
db4	94,49 %	89,76 %	76,38 %	10	80,52%
db8	95,42 %	84,97 %	81,70 %	10	73,37%
db20	97,53 %	86,42 %	82,30 %	10	72,00%

IV. KESIMPULAN

- Telah dilakukan Klasifikasi elektronik e-nose menggunakan untuk denoising dan kompresi dengan menggunakan Daubechie(db4, db8 dan db 20) dengan menggunakan dekomposisi di level 3 dan 6
- Hasil yang Kompresi dari dekomposisi sinyal *e-nose* retensi energi sinyal dengan rata-rata tingkat kesalahan rekonstruksi sinyal sebesar $3,57e^{-10}$, dengan nilai kesalahan terkecil $1,31e^{-12}$ dan kesalahan terbesar $1,62e^{-09}$
- Perubahan voltase terbesar ada pada sensor TGS822, dimana sensor TGS822 memiliki kepekaan terhadap gugus alkohol.
- Akurasi SVM mencapai 97,53% dalam mengenali pola-pola data e-nose hasil ekstrasi wavelet dan PCA.

REFERENCES

- Ricardo Gutierrez-Osuna, 2002. Pattern Analysis for Machine Olfaction: A Review. *IEEE Sensors Journal*, Vol.2, NO.3:189-201
- Krishna Persaud & George Dodd, 1982. Analysis of Discrimination Mechanisms in The Mammalian Olfactory System Using a Model Nose. Nature Publishing Group.
- B Moy, L., Tan, T. & Gardnet, J.W, 1994. Monitoring The Stability of Perfume and Body Odours with an Electronic Nose. *Perfumer and Flavourist*, 19, 11:19.
- Phaisangittisagul, Ekachai, 2007. Signal Processing Using Wavelets For Enhancing Electronic Nose Performance. *Electrical Engineering Faculty Of North Carolina State University*.
- Lav, R.K., 2006. Wavelet Transform: A Tool for Pattern Recognition of Olfactory Signal. *The Canadian Society for Bioengineering*
- Buck, L. and Axel, R, *Odorant Receptors and The Organization of The Olfactory System*. The Nobel Prize in Physiology or Medicine, 2004.

- H. Zwaardemaker and F. Hogewind, On spray-electricity and waterfall-electricity, Proc. Acad. Sci. Amst., 22, pp. 429-437, 1920.
- J. D. Hartman, A Possible Objective Method For The Rapid Estimation Of Flavors In Vegetables, Proc. Am. Soc. Hort. Sci., 64, pp. 335, 1954.
- J. W. Gardner and P. N. Bartlett, *Electronic Noses: Principles and Applications*. Oxford: Oxford University Press, 1999.
- M. Misiti, Y. Misiti, G. Oppenheim, and J-M. Poggi.2004. Wavelet toolbox user's guide: for use with Matlab. *MathWorks*.
- C. S. Burrus, R. A. Gopinath, and H. Guo.1998. Introduction to wavelets and wavelet transforms: a primer. *NJ, Prentice Hall*.
- Kotsiantis, SB 2007, Supervised machine learning: A review of classification techniques, Informatica, vol.31, no. 3, pp. 249-268.
- Smith, L. I., 2002. A tutorial on principal components analysis. URL http://www.cs.otago.ac.nz/cosc453/student_tutorials/principal_components.pdf. (URL diakses tanggal 6 Juni 2012).
- Vapnik V.N., 1999. The Nature of Statistical Learning Theory 2nd edition. Springererlag, New York Berlin Heidelberg
- Triyana, K., 1997. Prototipe metoda indentifikasi keluaran sistem sensor penciuman dengan jaringan neural artifisial, Tesis Megister, Pascasarjana ITB, Bandung.
- Yamanishi, T., 1995, Productio, processing, acceptance, nutrition and health: spesial issue on tea, Food Revies International, New York.